working paper
department
of economics

REPEATED GAMES WITH LONG-RUN
AND SHORT-RUN PLAYERS

By

Drew Fudenberg
David Kreps
Eric Maskin

No. 474   January 1988

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

REPEATED GAMES WITH LONG-RUN
AND SHORT-RUN PLAYERS
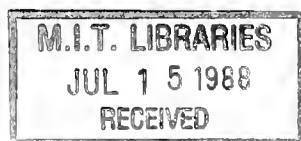
By

Drew Fudenberg
David Kreps
Eric Maskin

No. 474   January 1988

M.I.T., Department of Economics; Stanford University, Graduate School of
Business; and Harvard University, Department of Economics, respectively.

ABTRACT

This paper studies the set of equilibrium payoffs in games with long-
and short-run players and little discounting. Because the short-run players
are unconcerned about the future, equilibrium outcomes must always lie on
their static reaction (best response) curves. The obvious extension of the
Folk Theorem to games with this constraint would simply include the constraint
in the definitions of the feasible payoffs and of the minmax values. This
extension does obtain under the assumption that each player's choice of a
mixed strategy for the stage game is publicly observable, but, in contrast to
standard repeated games, the limit value of the set of equilibrium payoffs is
different if players can observe only their opponents' realized actions.

1. Introduction

The "folk theorem" for repeated games with discounting says that (under mild conditions) each individually-rational payoff can be attained in a perfect equilibrium for a range of discount factors close to one. It has long been realized that results similar to the folk theorem can arise if some of the players play the constituent game infinitely often and others play the constituent game only once, so long as all of the players are aware of all previous play. A standard example is the infinitely-repeated version of Selten's [1977] chain-store game, where a single incumbent faces an infinite sequence of short-run entrants in the game depicted in Figure 1. Each entrant cares only about its one-period payoff, while the incumbent maximizes its net present value. For discount factors close to one there is a perfect equilibrium in which entry never occurs, even though this is not a perfect equilibrium if the game is played only once or even a fixed finite number of times. In this equilibrium, each entrant's strategy is "stay out if the incumbent has fought all previous entry; otherwise, enter;" and the incumbent's strategies is "fight each entry as long as entry has always been fought in the past, otherwise acquiesce." Other examples of games with long and short run players are the papers of Dybvig-Spatt [1980] and Shapiro [1982] on a firm's reputation for producing high-quality goods and the papers of Simon [1951] and Kreps [1984] on the nature of the employment relationship.

This paper studies the set of equilibrium payoffs in games with long- and short-run players and little discounting. This set differs from what it would be if all players were long-run, as demonstrated by the prisoner's dilemma with one enduring player facing a sequence of short-run opponents. Because the short-run players will fink in every period, the only equilibrium

is the static one, no matter what the discount factor. In general, because the short-run players are unconcerned about the future, equilibrium outcomes must always lie on their static reaction (best response) curves. This is also true off of the equilibrium path, so the reservation values of the long-run players are higher when some of their opponents are short-run, because their punishments must be drawn from a smaller set.

The perfect folk theorem for discounted repeated games (Fudenberg-Maskin [1986]) shows that, under a mild full-dimensionality condition, any feasible payoffs that give all players more than their minmax values can be attained by a perfect equilibrium if the discount factor is near enough to one.[1] The obvious extension of this result to games with the constraint that short-run players always play static best responses would simply include that constraint in the definitions of the feasible payoffs and of the minmax values. Propositions 1 and 2 of Section 2 shows that this extension does obtain under the assumption that each player's choice of a mixed strategy for the stage game is publicly observable.

We then turn to the more realistic case in which players observe only their opponents' realized actions and not their opponents' mixed strategies. While in standard repeated games the folk theorem obtains in either case, when there are some short-run players the set of equilibria can be strictly smaller if mixed strategies are not observed. The explanation for this difference is that in ordinary repeated games, while mixed strategies may be needed during punishment phases, they are not necessary along the equilibrium path. In contrast, with short-run players some best responses, and thus some of the feasible payoffs, can only be obtained if the long-run players use mixed strategies. If the mixed strategies are not observable, inducing the long-run players to randomize may require that "punishments" occur with positive

probability even if no player has deviated. For this reason the set of equilibrium payoffs may be bounded away from the frontier of the feasible set.

Proposition 3 of Section 3 provides a complete characterization of the limiting value of the set of equilibrium payoffs for a single long-run player. This characterization, and the results of Section 2, assume that players have access to a publicly observable randomizing device. The device is used to implement strategies of the form: if player i deviates, then players jointly switch to a "punishment equilibrium" with some probability $p < 1$. While the assumption of public randomizations is not implausible, it is interesting to know whether it leads to a larger limit set of equilibrium payoffs. Proposition 4 in Section 4 shows that it does not: We construct "target strategies" in which a player is punished with probability one whenever his discounted payoff to date exceeds a target value, and shows that these strategies can be used to obtain as an equilibrium any of the equilibrium payoffs that were obtained via public randomizations in Proposition 3.

Proposition 3 shows that not all feasible payoffs can be obtained as equilibrium, so in particular we know that some payoffs cannot be obtained with the target strategies of Proposition 4. Inspection of that construction shows that it fails for payoffs that are higher than what the long-run player can obtain with probability one given the incentive constraint of the short-run players: For payoffs this high, there is a positive probability that player 1 will suffer a run of "bad luck" after which no possible sequence of payoffs could draw his discounted normalized value up to the target. As this problem does not arise under the criterion of time-average payoffs, one might wonder if the set of equilibrium payoffs is larger under time-averaging. Proposition 5 shows that the answer is yes. In fact, any feasible incentive compatible payoffs can arise as equilibria with time-averaging, so that we

3

obtain the same set of payoffs as in the case where the player's privately

mixed strategies are observable. This discontinuity of the equilibrium set in

passing from discounting to time averaging is reminiscent of a similar

discontinuity that has been established for the equilibria of repeated

partnership games (Radner [1986], Radner-Myerson-Maskin [1986]). The

relationship between the two models is discussed further in Section 5.

We have not solved the case of several long-run players and unobservable

mixed strategies. Section 5 gives an indication of the additional

complications that this case presents.

2.   Observable Mixed Strategies

Consider a finite n-player game in normal form,

$$g: \quad S_1 x...x S_n \rightarrow R^n \ .$$

We denote player i's mixed strategies by $\sigma \in \Sigma$, and write $g(\sigma)$ for the

expected value of g under distribution σ.

In this section we assume that a player can observe the others' past

mixed strategies. This assumption (or a restriction to pure strategies) is

standard in the repeated games literature, but as Fudenberg-Maskin [1986]

[1987a] have shown, it is not necessary there. (Here it matters -- see the

next section!) We will also assume that the players can make their actions

contingent on the outcome of a publicly observable randoming device.

Label the players so that players 1 to j are long-run and j+1 to n are

short-run. Let

$$B: \quad \Sigma_1 x...x \Sigma_j \overrightarrow{\rightarrow} \Sigma_{j+1}...x \Sigma_n$$

4

be the correspondence which maps any strategy selection $(\sigma_1, \ldots, \sigma_j)$ for the long-run players to the corresponding Nash equilibria strategy selections for the short-run players. If there is only one short-run player, $B(\sigma)$ is his best response correspondence.

For each i from 1 to j, choose $m^i = (m^i_1, \ldots, m^i_n)$ so that m^i solves

$$\min_{m^i \in \text{graph}(B)} \ \max_{\sigma_i} \ g^i(\sigma_i, m^i_{-i}),$$

and set

$$\underline{v}_i = \max_{\sigma_i} g^i(\sigma_i, m^i_{-i}).$$

(This minimum is attained because the constraint set graph(B) is compact and the function $\max_{\sigma_i} g^i(\sigma_i, m^i_{-i})$ is continuous in m^i_{-i}.)

The strategies m^i_{-i} minimize long-run player i's maximum attainable payoff over the graph of B. The restriction to this set reflects the constraint that the short-run players will always choose actions that are short-run optimal. Given this constraint, no equilibrium of the repeated game can give player i less than $\underline{v}_i$. (In general, the short-run players could force player one's payoff even lower using strategies that are not short-run optimal). Note that m^i specifies player i's strategy m^i_i, which need not be a best response to m^i_{-i}: Player i must play in a certain way to induce the short-run players to attain the minimum in the definition of m^i. In order to construct equilibria in which player i's payoff is close to $\underline{v}_i$, player i will need to be provided with an incentive to cooperate in his own punishment.

In the repeated version of g, we suppose that long-run players maximize the discounted normalized sum of their single-period payoffs, with common discount factor δ. That is, long-run player i's payoff is $(1-\delta)\sum_{t=0}^{t=\infty} \delta^t g_i(\sigma(t))$. Short-run players in each period act to maximize that period's payoff. All players, both long- and short-run, can condition their play on all previous actions.

Let $U = \{ v = (v_1, \ldots v_n) \mid \exists \sigma \text{ in graph (B) with } g(\sigma) = v\}$

Let V = convex hull of U;

and let $V^* = \{v \in V \mid \text{for all i from 1 to j, } v > \underline{v}_i\}$.

We call payoffs in V^* attainable payoffs for the long-run player. Only payoffs in V* can arise in equilibrium. We begin with the case of a single long-run player.

Proposition 1: If only player one is a long-run player, then for any $v_1 \in V^*$ there exists a $\underline{\delta} \in (0,1)$ such that for all $\delta \in (\underline{\delta}, 1)$, there is a subgame-perfect equilibrium of the infinitely repeated game with discount factor δ in which player i's discounted normalized payoff is v_i.

Proof: Fix a $v_i \in V^*$ and consider the following strategies. Begin in Phase A, where players play a $\sigma \in$ graph (B) (or a public randomization over such σ's) that gives player 1 payoff v_1. Deviations by the short-run players are ignored. If player one deviates, he is punished by players switching to the punishment strategy m^1 for $T(\delta)$ periods, after which play returns to Phase A; if $T(\delta)$ is large enough, deviations in Phase A are unprofitable. Now m_1^1 need not be a best response against m_{-1}^1, so we must insure that player one

does not prefer to deviate during the punishment phase. This is done by specifying that a deviation in this phase restarts the punishment. Since the most that player 1 can obtain in any period of the punishment phase is $\underline{v}_1$, he will prefer not to deviate so long as $T(\delta)$ is short enough that player 1's normalized payoff at the start of the punishment phase is at least $\underline{v}_1$. Let $\bar{v}_1$ = $\max\limits_{\sigma \in \text{graph}(B)} g_1(\sigma)$. The two constraints on $T(\delta)$ will be satisfied if:

(1) $\quad (1-\delta)\bar{v}_1 + \delta(1-\delta^{T(\delta)})g_1(m^1) + \delta^{T(\delta)+1}v_1 \le v_1$, or equivalently

(1') $\quad \delta^{T(\delta)+1} \le (v_1 - \delta g_1(m^1) + (1-\delta)\bar{v}_1)/(v_1 - g_1(m^1))$, and

(2) $\quad (1-\delta^{T(\delta)})g_1(m^1) + \delta^{T(\delta)}v_1 > \underline{v}_1$, or equivalently

(2') $\quad \delta^{T(\delta)} > (\underline{v}_1 - g_1(m^1))/(v_1 - g_1(m^1))$.

The right-hand sides of inequalities (1') and (2') have the same denominator, and for δ close to 1 the numerator of (1') exceeds the numerator of (2'). Then since δ^T is approximately continuous in T for δ close to 1, we can find a $\underline{\delta} < 1$ such that for all greater δ there is a $T(\delta)$ satisfying (1') and (2').

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ Q.E.D.

In repeated games with three or more players, a full-dimensionality condition is required for all feasible individually rational payoffs to be enforceable when δ is near enough to one. The corresponding condition here is that the dimensionality of V* equals the number of long-run players.

Proposition 2: Assume that the dimensionality of V* = j, the number of long-run players. Then for each v in V*, there is a $\underline{\delta} \in (0,1)$ such that for all $\delta \in (\underline{\delta},1)$ there is subgame-perfect equilibrium of the infinitely repeated game with discount factor δ in which player i's normalized payoff is v_i.

Remark: The proof of Proposition 2 follows that of Fudenberg-Maskin's Theorem 2: If a (long-run) player deviates, he is punished long enough to wipe out the gain from deviation. To induce the other (long-run) players to punish him, they are given a "reward" at the end of the punishment phase. One small complication not present in Fudenberg-Maskin is that, as in Proposition 1, the player being punished must take an active role in his punishment. This, however, can be arranged with essentially the same strategies as before.

Proof: Choose a σ (or a public randomization over several σ's) so that $g(\sigma) = v$. Also choose v' in the interior of V* and an $\varepsilon > 0$ so that for all i from 1 to j $(v_1'+\varepsilon,\ldots,v_{i-1}'+\varepsilon,v_i',v_i'+\varepsilon,\ldots,v_j'+\varepsilon)$ is in V* and $v_i' + \varepsilon < v_i$. Let T^i be a joint strategy that yields $v_j' + \varepsilon$ to all the long-run players but i, and yields v_i' to i. Let $w_i^j = g_i(m^j)$ be player i's period payoff when j is being punished with the strategies m^j. For each i, choose an integer N_i so that

$$\bar{v}_i + N_i \underline{v}_i < (N+1)\, v_i'$$

where $\bar{v}_i = \max g^i$ is i's greatest one-period payoff.

Consider the following repeated-game strategy for player i:

(0)   Obey the following rules regardless of how the short-run players have played in the past:

(A)   Play σ_i each period as long as all long-run players played σ last period, or if σ had been played until last period and two or more long-run players failed to play σ last period.

If long-run player j deviates from (A), then

(B_j)   Play m_i^j for N_j periods, and then

(C)   Play T_i^j thereafter.

If long-run player k deviates in phase (B_j) or (C), then begin phase (B_j) again with j = k. (As in phase A, players ignore simultaneous deviations by two or more long-run players.)

As usual, it suffices to check that in every subgame no player can gain from deviating once and then conforming. The condition on N_i ensures that for δ close to one, the gain from deviating in Phase A or Phase C is outweighed by Phase B's punishment. If player j conforms in B_j (i.e. when she is being punished) her payoff is at least $q_j \equiv (1-\delta^{N_j})w_j^j + \delta^{N_j}v_j'$, which exceed $\underline{v}_j$ if δ is close enough to one. If she deviates once and then conforms, she receives at most $\underline{v}_j$ the period she deviates, and postpones the payoff $q_j > \underline{v}_j$, which lowers her payoff. If player k deviates in Phase B_j, she is minmaxed for the next N_k periods and Phase-C play will give her v_k' instead of $v_k' + \varepsilon$. Thus it is easy to show that such a deviation is unprofitable. (See Fudenberg-Maskin for the missing computations.)

3.   Unobservable Mixed Strategies

We now drop the assumption that players can observe their opponents'
mixed strategies, and instead assume they can only observe their opponents'
realized actions. In ordinary repeated games, (privately) mixed strategies
are needed during punishment phases, because in general a player's minmax
value is lower when his opponents use mixed strategies. However, mixed
strategies are not required along the equilibrium path, since desired play
along the path can be enforced by the threat of future punishments.
Fudenberg-Maskin showed that, under the full-dimension condition of
Proposition 2, players can be induced to use mixed strategies as punishments
by making the continuation payoffs at the end of a punishment phase dependent
on the realized actions in that phase in such a way that each action in the
support of the mixed strategy yields the same overall payoff.

In contrast, with short-run players some payoffs (in the graph of B) can
only be obtained if the long-run players privately randomize, so that mixed
strategies are in general required along the equilibrium path. As a
consequence, the set of equilibrium payoffs in the repeated game can be
strictly smaller when mixed strategies are not observable. This is
illustrated by the following example of a game with one long-run player, Row,
and one short-run player, Col.

Let p be the probability that Row plays D. Col's best response is M if
$0 \leq p \leq 1/2$, L if $1/2 \leq p \leq 100/101$, and R if $p \geq 100/101$. There are three
static equilibria: the pure strategy equilibrium (D,R), a second in which $p =$
$1/2$ and Col mixes between M and L, and a third in which $p = 100/101$ and Col
mixes between L and R. Row's maximum attainable payoff is 3, which occurs
when $p = 1/2$ and Col plays L.

	L	M	R
U	4, 0	0, 1	-1, -100
D	2, 2	1, 1	0, 3

Figure 1

If Row's mixed strategy is observable, she can attain this payoff in the infinitely repeated game if δ is near enough to 1. If however Row's mixed strategy is not observable, her highest equilibrium payoff is at most 2 regardless of δ.

To see this, fix a discount factor δ, and let $v^*(\delta)$ be the supremum over all Nash equilibria of ROW's equilibrium payoff. Suppose that for some δ $v^*(\delta) = 2 + \varepsilon' > 2$, and choose an equilibrium σ such that player 1's payoff is $v(\sigma) = v^*(\delta) - \varepsilon > 2$. It is easy to see that the set of equilibrium payoffs is stationary: Any equilibrium payoff is an equilibrium payoff for any subgame, and conversely. Thus, the highest payoff player 1 can obtain starting from period 2 is also bounded by $v^*(\delta)$. Since $v(\sigma)$ is the weighted average of player 1's first-period payoff and her expected continuation payoff, player 1's first-period payoff must be $v^*(\delta) - \varepsilon/(1-\delta)$. For ε sufficiently small, this implies that player 1's first period payoff must exceed 2.

In order for Row's first-period payoff to be at least 2, Col must play L with positive probability in the first period. As Col will only play L if Row randomizes between U and D, Row must be indifferent between her first period choices, and in particular must be willing to play D. Let v_D be Row's expected payoff from period 2 on if she plays D in the first period. Then we

must have

(3) $2(1-\delta) + \delta v_D = v\star.$

But since $v_D \leq v\star$, we conclude that $v\star \leq 2$.

While Row cannot do as well as if her mixed strategies were observable, she can still gain by using mixed strategies. For δ near enough to one there is an equilibrium which gives Row an normalized payoff of 2, while Row's best payoff when restricted to pure strategies is the static equilibrium yielding 1. To induce Row to mix between U and D, specify that following periods when Col expects mixing and Row plays U, play switches with probability p to (D,R) for ten periods and then reverts to Row randomizing and Col playing L. The probability p is chosen so that Row is just indifferent between receiving 2 for the next eleven periods, or receiving 4 today and risking punishment with probability p. This construction works quite generally, as shown in the following proposition.

Proposition 3: Consider a game with a single long-lived player, player one, and let

$$v_1^\star = \max_{\sigma \in \text{graph } B} \quad \min_{s_1 \in \text{supp } \sigma_1} \quad g_1(s_1, \sigma_{-1}).$$

Then for any $v_1 \in (\underline{v}_1, v_1^\star)$ there exists a $\delta' < 1$ such that for all $\delta \in (\delta', 1)$, there is an equilibrium in which player one's normalized payoff is v_1. For no δ is there an equilibrium where player one's payoff exceeds $v_1^\star$

<u>Proof</u>: We begin by constructing a "punishment equilibrium" in which player one's normalized payoff is exactly $\underline{v}_1$. If $\underline{v}_1$ is player one's payoff in a static equilibrium this is immediate, so assume all the static equilibria given player one more than $\underline{v}_1$. The strategies we will use have two phases. The game begins in phase A, where the players use m^1, a strategy which holds player one's maximum one-period payoff to $\underline{v}_1$. If player one plays s_1, players publicly randomize between remaining in phase A and switching to a static Nash equilibrium for the remainder of the game. If e_1 is one's payoff in this static equilibrium, set the probability of switching after x_1, $p(s_1)$, to be

$$
(4) \qquad p(s_1) = \frac{(1-\delta)(\underline{v}_1 - g_1(s_1, m^1_{-1}))}{\delta(e_1 - g_1(s_1, m^1_{-1}))} .
$$

(If δ is near enough to one, $p(s_1)$ is between 0 and 1.)

The switching probability has been constructed so that player one's normalized profit is $\underline{v}_1$ for all actions, including those in the support of m^1_1, so she is indifferent between these actions.

Next we construct strategies yielding v^*_1 for $v^*_1 > \underline{v}_1$. Let $\sigma^* = (\sigma^*_1, \sigma^*_{-1})$ be the corresponding mixed strategies. Play begins in phase A with players following a. If player one deviates to an action outside the support of σ_1, then switch to the "punishment equilibrium" constructed above. If player one plays an action s_1 in the support of σ_1, then switch to the punishment equilibrium with probability $p(s_1)$, and otherwise remain in phase A. The probability $p(s_1)$ is chosen so that player one's payoff to all actions in the support of σ_1 is v^*_1. As above, this probability exists if δ is near enough to one. These strategies are clearly an equilibrium for large δ.

Equilibrium payoffs between $\underline{v}_1$ and v_1^* are obtained by using public randomizations between those two value. The argument that player one's payoff cannot exceed v_1^* is exactly as in the example.

4. No Public Randomizations

The equilibria that we constructed in the proofs of Theorems 1 through 3 relied on our assumption that players can condition their play on the outcome of a publicly observed random variable. While that assumption is not · implausible, it is also of interest to know whether the assumption is necessary for our results. For this reason, Proposition 4 below extends Proposition 3 to games without public randomizations. (We have not thought about the possible extension of Propositions 1 and 2 because we think the situation without public randomizations but where private randomization can be verified ex-post is without interest.) The intuition, as explained in Fudenberg-Maskin [1987c], is that public randomizations serve to convexify the set of attainable payoffs, and when δ is near to 1 this convexification can be achieved by sequences of play which vary over time in the appropriate way. Fudenberg-Maskin [1987c] shows that public randomizations are not necessary for the proof of the perfect Folk Theorem. However, as we have already seen, there are important differences between classic repeated games and repeated games with some short-run players, so the fact that public randomizations are not needed for the folk theorem should not be thought to settle the question here.

Proposition 4: Consider a game with a single long-run player, player 1, where public randomizations are not available. As in Proposition 3, let

$$v_1^* = \max_{\sigma \in \text{graph}(B)} \min_{s_1 \in \text{supp } \sigma_1} g_1(s_1, \sigma_{-1}),$$

and let σ^* be a strategy that attains this max. Then, for any $v_1 \in (\underline{v}_1, v_1^*)$ there exists a $\delta' < 1$ such that for all $\delta \in (\delta', 1)$ there is a subgame-perfect equilibrium where player 1's discounted normalized payoff is v_1.

Remark: Fix a static Nash equilibrium $\hat{\sigma}$ with payoffs $\hat{v}$. For each v_1 the proof constructs strategies that keep track of the agent's total realized payoff to date t and compares it to the "target" value of $(1-\delta^t) v_1$, which is what the payoff to date would be if the agent received v_1 in every period. If v_1 exceeds the payoff in a static equilibrium, then play initially follows the (possibly mixed) strategy σ^*, and whenever the realized total is sufficiently greater than the target value, the agent is "punished" by reversion to the static equilibrium. If the target is less than the static equilibrium, then play starts out at the (possibly mixed) strategy m^1, with intermittent "rewards" of the static equilibrium whenever the realized payoff drops too low.

Proof:

(A) It is trivial to obtain $\hat{v}_1$ as an equilibrium payoff.

(B)   To attain any payoff v_1 between $\hat{v}_1$ and v_1^* we proceed as follows. Renormalize the payoffs so that $\hat{v}_1 = 0$, and take δ large enough that $(1-\delta)\bar{v}_1 < v_1^*$. Define $J_0 = 0$ and $\delta^*(0) = \sigma^*$, and for each time t>0 define the strategies $\delta^*(h_t)$ and an index J_t as follows:

$$J_t \equiv J_{t-1} + (1-\delta)\delta^{(t-1)} g_1(s_1(t-1), \delta_{-1}^*(h_{t-1}),$$

where $s_1(t-1)$ is player 1's action in period t-1 (as opposed to his choice of mixed strategy,) and $R_t \equiv (1-\delta^t)v_1$. If player 1's payoff were v_1 each period, then his accrued payoff J_t would equal R_t. The equilibrium strategies will "punish" the agent whenever J_t exceeds R_t by too large a margin. More precisely, we define

$$\delta^*(h_t) = \begin{cases} \hat{\sigma} & \text{if } J_t \geq R_{t+1} \text{ and } J_\tau \geq R_\tau \text{ for all } \tau \leq t \\ \sigma^* & \text{if } J_t < R_{t+1} \text{ and } J_\tau \geq R_\tau \text{ for all } \tau \leq t \\ \hat{\sigma} & \text{if } J_\tau < R_\tau \text{ for any } \tau \leq t \end{cases}$$

Note that since J_t is a discounted sum, for each infinite history h_∞, J_t converges to a limit J_∞. Moreover, as long as the other players use strategy δ_{-1}^*, player 1's payoff to any strategy is simply the expected value of J_∞, and his expected payoff in any subgame starting at time t is $\delta^{-t}(J_\infty - J_t)$.

    We will now argue that (i) if player 1 uses strategy δ_1^*, then $J_t \geq R_t$ for all times t and histories h_t, which imples that $J_\infty \geq v_1$, (ii) that regardless of how player 1 plays, $J_\infty \leq v_1$, so player 1's payoff in the subgame starting at time t is bounded by $\delta^{-t}(v_1 - J_t)$ for all histories h_∞, and (iii) that in any subgame where at some $\tau \leq t$ $J_\tau < R_\tau$, it is a best response for all players to follow the prescribed strategy of always playing

the static equilibrium $\hat{\sigma}$, and so $J_\infty < v_1$.

Conditions (i) and (ii) imply that it is a best response for player 1 to play ∂^* in every subgame where J_τ has never dropped below R_τ, and that player 1's equilibrium payoff is v_1. Condition (iii), whose proof is immediate, says that following ∂^* is also a Nash equilibrium in subgames where J_τ has dropped below R_τ, so that ∂^* is a subgame-perfect equilibrium. (The condition that the short-run players not wish to deviate is incorporated in the construction of ∂^*.)

Proof of (i): We must show that if player 1 follows ∂_1^* then for all t, $J_t \geq (1-\delta^t)v_1 = R_t$. Since $J_0 = 0$, this is true for t=0. Assume it is true for t=τ. At period τ, either (a) $J_\tau \geq R_{\tau+1}$ or (b) $J_\tau < R_{\tau+1}$. In case (a), $\partial^*(t) = \hat{\sigma}$. Since $g_1(s_1, \hat{\sigma}_{-1}) = 0$ for every pure strategy s_1 in the support of $\hat{\sigma}_1$, we have $J_{\tau+1} = J_\tau$, and $J_\tau \geq R_\tau$ by inductive hypothesis. In case (b), $\partial^*(h_\tau) = \sigma^*$, so $\min \{g_1(s_1(\tau), \partial_{-1}^*(h_\tau)) \mid s_1(\tau) \in \text{support}(\partial_1^*(h_t))\} = v_1^*$,

and $J_{\tau+1} \geq J_\tau + (1-\delta)\delta^\tau v_1^* \geq (1-\delta^\tau)v_1^* + (1-\delta)\delta^\tau v_1^* = (1-\delta^{\tau+1})v_1 = R_{\tau+1}$,

where the second inequality comes from the inductrive hypothesis. Thus $J_t \geq (1-\delta^t)v_1^*$ for all t, and so if player 1 follows ∂_1^* then $J_\infty \geq v_1$.

Proof of (ii): Next we claim that regardless of how player 1 plays, $J_\infty \leq v_1$. If for some history $J_\infty > v_1$, then there is a T' such that $J_t \geq (1-\delta^{t+1})v_1$ for all $t > T'$. Thus $\partial_{-1}^*(h_t) = \hat{\sigma}$ for all $t > T'$, and since the most player 1 can get when his opponents play $\hat{\sigma}$ is zero, we have that $J_\infty \leq J_{T'}$. Let T be the smallest such T' so that $J_{T'-1} < (1-\delta^T)v_1$. (Since $J_0 = 0$, $T > 0$.) Then $J_\infty \leq J_T < J_{T-1} + \delta^T(1-\delta)\bar{v}_1 < J_{T-1} + \delta^T v_1 < v_1$, where we have substituted in our bound on δ. This argument also shows that player 1's payoff in the subgame starting at time t is bounded by $\delta^{-t}(v_1 - J_t)$, and from part (i) this

17

payoff can be attained by following $\hat{\sigma}_1^*$.

(C): Next we show how to construct equilibria (for large enough δ) that yield payoffs v_1 between $\underline{v}_1$ and the static equilibrium payoff of zero. Pick a $v_1 \in (\underline{v}_1, 0)$, and choose δ large enough that $(1-\delta)\min_{\sigma} g_1(\sigma) > v_1$. Then set $J_0 = 0$, and $\hat{\sigma}^*(0) = m^1$. Now define $J_t(h_t)$ and $\hat{\sigma}(h_t)$ as follows. Set $J_t = J_{t-1} + \delta^t g_1(s_1(t), \hat{\sigma}_{-1}^*(h_t))$, and set

$$\hat{\sigma}^*(h_t) = \begin{cases} \hat{\sigma} & \text{if } J_t < R_{t+1} \\ \\ m^1 & \text{if } J_t \geq R_t \end{cases}$$

Proceeding as above, we claim that

(i) if player 1 uses strategy $\hat{\sigma}_1^*$, then $J_t \geq R_t$
for all times t and histories h_t, which imples that $J_\infty \geq v_1$,

(ii) that regardless of how player 1 plays, $J_\infty \leq v_1$, so player 1's payoff in the subgame starting at time t is no greater than $\delta^{-t}(v_1 - J_t)$, and

(iii) that in subgames where $J_t < R_t$ it is a best response for player one to play $\hat{\sigma}^*(h_t) = \hat{\sigma}_1$.

Proof of (i): We must show that if player 1 follows $\hat{\sigma}_1^*$ then for all t, $J_t \geq (1-\delta^t)v_1 = R_t$. Since $J_0 = 0$, this is true for t=0. Assume it is true for $t=\tau$. At period τ, either (a) $J_\tau \geq R_{\tau+1}$ or (b) $J_\tau < R_{\tau+1}$.

In case (a),

$$J_{\tau+1} \geq J_\tau + \delta^\tau (1-\delta)\min(g_1) \geq J_\tau + \delta^\tau v_1 \text{ (from our bound on } \delta)$$

$$\geq R_\tau + \delta^\tau v_1 \text{ (by inductive hypothesis)}$$

$$= R_{\tau+1}.$$

In case (b), $\hat{\partial}^*(h_\tau)=\hat{\sigma}$, so $g_1(s_1(\tau), \hat{\partial}^*_{-1}(h_\tau))=0$ for all $s_1(\tau) \in$ support $\hat{\partial}^*_1(h_\tau)$,

and $J_{\tau+1} = J_\tau \geq R_\tau \geq R_{\tau+1}.$

Thus $J_t \geq (1-\delta^t)v_1^*$ for all t, and so if player 1 follows $\hat{\partial}^*_1$ then $J_\infty \geq v_1$.

Proof of (ii): We claim that for all strategies of player 1 and all times t and histories h_t, $J_t \leq (1-\delta^t)v_1 = R_t$. Since $J_0 = 0$, this is true for t=0. Assume it is true for $t=\tau$. Then at period τ, either (a) $J_\tau < R_{\tau+1}$ or (b) $J_\tau \geq R_{\tau+1}$. In case (a), $\hat{\partial}^*(t) = \hat{\sigma}$, so $J_{\tau+1} \leq J_\tau < R_{\tau+1}$.
In case (b), $\hat{\partial}^*(h_\tau) = m^1$, so $\max_{s_1} g_1(s_1, \hat{\partial}^*_{-1}(h_\tau)) = \underline{v}_1$, and

$$J_{\tau+1} < J_\tau + (1-\delta)\delta^\tau \underline{v}_1 < (1-\delta^\tau)v_1 + (1-\delta)\delta^\tau \underline{v}_1 \text{ (by the inductive hypothesis)}$$

$$\leq (1-\delta^{\tau+1})v_1 \text{ (from the bound on } \delta) = R_{\tau+1}.$$

Thus $J_t < (1-\delta^t)v_1$ for all t, and so regardless of how player 1 plays, $J_\infty \leq v_1$.

(iii)   Conditions (i) and (ii) show that in any subgame with $J_t \geq R_t$ player 1 can attain the upper bound of v_1 by following $\hat{\partial}^*_1$. Now we consider subgames with $J_t < R_t$. If $J_t \leq v_1$, then regardless of how player 1 plays, we will have $J_\tau \leq R_\tau$ for all $\tau \geq t$, so player 1's opponents will play $\hat{\sigma}_1$ for the remainder of the game. Here it is clearly a best response for player 1 to play $\hat{\sigma}_1 = \hat{\partial}^*_1$. If $J_t > v_1$, then by playing $\hat{\sigma}_1$ player 1 can ensure that $J_\tau \geq R_\tau$ at some $\tau>t$, which ensures that player 1 attains a payoff of v_1 in the subgame starting at t. If player 1 instead chooses a strategy which

assigns positive probability to the event that $J_\tau < R_\tau$ for all $\tau > t$, he can only lowe his payoff: The payoff for histories with $J_\infty < R_\infty$ is less than v_1, and the payoff for the histories with $J_\infty \geq R_\infty$ is bounded above by v_1.

<div align="right">Q.E.D.</div>

Proposition 4 shows how to attain any payoffs between $\underline{v}_1$ and v_1^* by means of "target strategies." From Proposition 3 we know that such strategies cannot be used to attain higher payoffs. We think that it is interesting to note where an attempted proof would fail.

In part (A), we proved that if player 1 followed δ^* then for *every* sequence of realizations player 1's payoff is at least v_1. Imagine that we try to attain a payoff $v_1 > v_1^*$ by setting the target $R_t = (1-\delta^t)v_1$. Then in the "reward" phases where σ^* is played, it might be that player 1's realized payoff is less than v_1. (Recall that by definition it cannot be lower than v_1^*). After a sufficiently long sequence of these outcomes, player 1's realized payoff J_t would be so much lower than v_1 that even receiving the best possible payoff at every future date would not bring his discounted normalized payoff up to the target.

This problem of going so far below the target that a return is impossible does not arise with the criterion of time-average payoffs, since the outcomes in any finite number of periods are then irrelevant. For this reason we <u>can</u> attain payoffs above v_1^* under time averaging, as we show in the next section.

5. <u>Time Averaging</u>

The reason that player 1's payoff is bounded by what he obtains when she plays her least favorite strategy in the support of σ^* is that every time she plays a different action she must be "punished" in a way that makes all of the actions in σ^* equally attractive. A similar need for "punishments" along the

equilibrium path occurs in repeated partnership games, where two players make an effort decision that is not observed by the other, and the link between effort and output is stochastic. Since shirking by either player increases the probability of low output, low output must provoke punishment, even though low output can occur when neither player shirks. This is why the best equilibrium outcome is bounded away from efficiency when the payoff criterion is the discounted normalized value. (Radner-Myerson-Maskin [1986], Fudenberg-Maskin [1987a]). However, Radner [1986] has shown that efficient payoffs can be attained in partnerships with time averaging. His proof constructed strategies so that (1) if players never cheat, punishment occurs only finitely often, and thus is negligible, and (2) an infinite number of deviations is very likely to trigger a substantial punishment. Since no finite number of deviations can increase the time-average payoff, in equilibrium no one cheats yet the punishment costs are negligible.

Since the inefficiencies in repeated partnerships and games with short-run players both stem from the need for punishments along the equilibrium path, it is not surprising that the inefficiencies in our model also disappear when players are completely patient. We prove this with a variant of the "target strategies" we used in Section 4. These strategies differ from Radner's in that even if player 1 plays the equilibrium strategy, she will be punished infinitely often with probability one. However, along the equilibrium path the frequency of punishment converges to zero, so that as in Radner the punishment imposes zero cost.

Proposition 5: Imagine that player 1 evaluates payoff streams with the criterion $\liminf_{T \to \infty} E \, (1/T) \sum_{t=0}^{t=T-1} g_1(s(t))$. Then for all $v_1 \in V^*$ there is a subgame-perfect equilibrium with payoffs v_1.

Remark: The proof is based on a strong law of large numbers for martingales with independent increments,[2/] which we extend to cover the difference between a supermartingale and its lowest value to date. The relevant limit theory is developed in the Appendix.

Proof: As in Proposition 4, we use different strategies for payoffs above and below some fixed static equilibrium $\hat{\sigma}$. Imagine that v_1 exceeds player 1's payoff in this equilibrium, and normalize $v_1 = 0$. Let $\grave{\sigma}$ be the (possibly mixed) strategy in graph (B) that maximizes player one's expected payoff, and define $g_1(\grave{\sigma}) = \grave{v}_{1t=T}$

Define $J_0 = 0$ and $J_T = \sum_{t=0}^{T-1} g_1(s_1(t), s_{-1}(t))$. (This differs from the definition in Section 4, where we used player 1's realized action and the mixed strategy of her opponents in defining J_t). Note that player 1's objective function is lim inf $E(1/T)J_T$ Set

$$
\vartheta^*(h_t) = \begin{cases} \grave{\sigma} & \text{if } J_t \geq 0 \\ \hat{\sigma} & \text{if } J_t < 0 \end{cases}
$$

We claim that (i) no matter how player 1 plays, her payoff is bounded by v_1, and (ii) that by following s_1^* player 1 can attain payoff v_1 almost surely (and hence in expectation.)

To prove this, let $\vartheta(h_t)$ be an arbitrary strategy for player 1, and fix the associated probability distribution over infinite-horizon histories. For each history, let $R_t(h_t) = \{\tau \leq t-1 \mid \vartheta^*(h_\tau) = \grave{\sigma}\}$ be the "reward" periods, and let

$P_t(h_t) = \{\tau \le t-1 \mid \partial^*(h_\tau) = \sigma\}$ be the "punishment" ones.

Then let $M_t(h_t) = \sum_{\tau \in R_t} g_1(x(t))$ be the sum of player 1's payoffs in the good

periods, and set

$$N_t(h_t) = \sum_{\tau \in P_t} g_1(x(t)).$$

Note that the reward and punishment sets and the associated scores are defined

path wise, i.e. they depend on the history h_t; henceforth, though, we will

omit the history h_t from the notation. Finally define $\hat{M}_t = \max_{\tau \le t} M_t$, $\hat{N}_t = \min_{\tau \le t}$

N_τ, and $v_1 = \min g_1(\sigma)$

We claim that for all t,

(5) $\qquad v_1 + (M_t - \hat{M}_t) \le J_t \le \bar{v}_1 + (N_t - \hat{N}_t).$

This is clearly true for $t = 0$. Assume (5) holds for all $\tau \le t$. At the start

of period t, either (a) $J_t > 0$ or (b) $J_t \le 0$. In case (a), $J_{t+1} \ge v_1 + J_t \ge v_1 \ge$

$v_1 + (M_t - \hat{M}_t)$. Also, since $\partial^*(h_t) = \bar{\sigma}$,

$J_{t+1} = J_t + N_{t+1} - N_t \le \bar{v}_1 + N_{t+1} - \hat{N}_t \le \bar{v}_1 + (N_{t+1} - \hat{N}_{t+1})$, so that (5) is satisfied.

In case (b), $J_{t+1} \le \bar{v}_1 + J_t \le \bar{v}_1 \le \bar{v}_1 + (N_{t+1} - \hat{N}_{t+1})$,

and $J_{t+1} = J_t + M_{t+1} - M_t \ge v_1 + M_{t+1} - \hat{M}_t \ge v_1 + (M_{t+1} - \hat{M}_{t+1})$,

so once again (5) is satisfied.

Lemmas 3 in the Appendix shows that $(N_t-\hat{N}_t)/t$ converges to zero almost surely. Since the per-period payoffs are uniformly bounded, this implies that limsup $(1/T)J_T \leq 0$ almost surely, and since the per-period payoffs are uniformly bounded, limsup $1/T\ E(J_T) \leq 0$ as well. Lemma 4 shows that if player 1 plays so that M_t is a submartingale, then the $(M_t-\hat{M})$ converges to zero as well. Since this is true when player 1 follows s_1^*, the result follows.

<div align="right">Q.E.D.</div>

We can show that with our strategies, player 1 is punished infinitely often $(J_t > 0)$ with probability one. This contrasts with Radner's construction of efficient equilibria for symmetric time-average partnership games, where the probability of infinite punishment is zero. It seems likely that our "target-strategy" approach provides another way of constructing efficient equilibria for those games; it would be interesting to know whether this could be extended to asymmetric partnerships. Our approach has the benefit of making more clear why the construction cannot be extended to the discounting case. It also avoids the need to invoke the law of the iterated logarithm, which may make the proof more intuitive, although we must use the strong law for martingales in its place.

6. Several Long Run Players with Unobservable Mixed Strategies

The case of several long-run players is more complex, and we have not completely solved it. As before, we can construct mixed-strategy equilibria in which the long-run players do better than in any pure strategy equilibrium, and once again they cannot do as well as if their mixed strategies were directly observable. However, we do not have a general characterization of the enforceable payoffs. Instead, we offer an example of payoffs that cannot

be enforced, and a very restrictive condition that suffices for enforceability.

Figure 2 presents a 3-player version of the game in Figure 1. Row's and Col's choices and payoffs are exactly as before. The third player, DUMMY, who is a long-run player, receives 3 if Col plays L and receives 0 otherwise. The feasible payoffs for Row and DUMMY are depicted in Figure 2. Consider the feasible point at which $p = 1/2$ and Col plays L. Here Row and Dummy both receive 3. The argument of Section 3 shows that Row's best equilibrium payoff is not 3 but 2, which is the minimum of payoff over the actions in the support of her mixed strategy. Dummy is not mixing, so Dummy's minimum payoff over the support of her strategy is 3. (Indeed this is the minimum over the support of the produce of the two strategies.) Thus one might hope that, by analogy to the proof of Proposition 3, we could show that the payoffs (2,3) were enforceable. But these payoffs are not even feasible! The highest Dummy's payoff can be when Row's payoff is 2 is $2 \frac{196}{205}$. (See Figure 3, which depicts the feasible set.) The problem is that an equilibrium in which Row usually randomizes must sometimes have Col play M or R to "tax away" Row's "excess gains" from playing U instead of D, and this "tax" imposes a cost on Dummy.

	L	R	M
U	4 , 3 , 0	0 , 0 , 1	-1 , 0 , -100
D	2 , 3 , 0	1 , 1 , 1	0 , 0 , 3

Player two is a "dummy", player three chooses COLs.

Figure 2

Player 2's payoff

Player 1's payoff
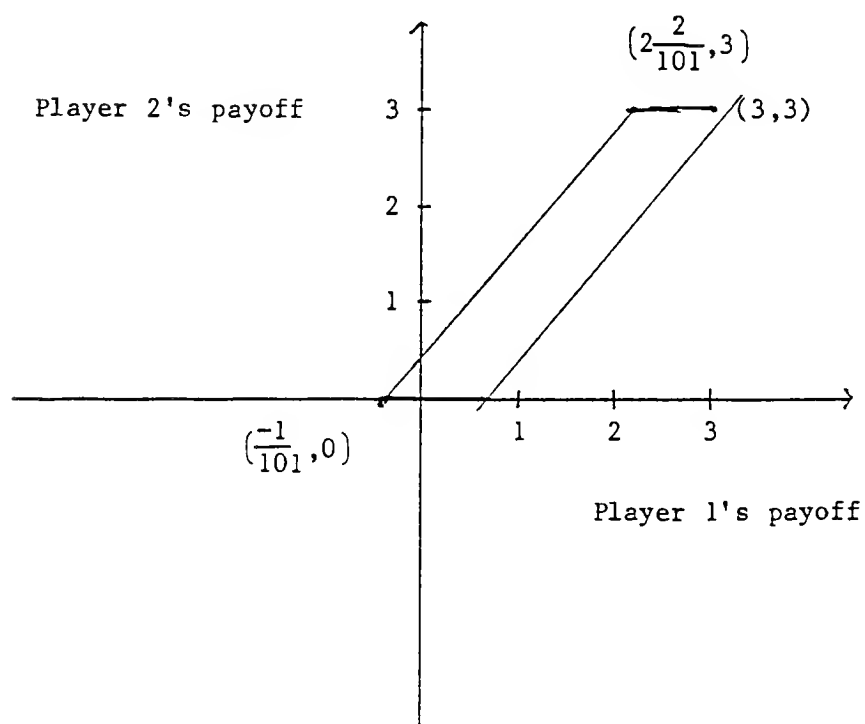
Feasible set when three plays a SR best response.

Figure 3

Next, consider the game in Figure 4.

	L	R
U	1, 1, 1	2, 0, 1
D	0, 2, 1	-1, -1, -9

A

	L	R
U	1, 1, -99	2, 0, 1
D	0, 2, 1	-1, -1, -1

B

	L	R
U	14, 4, 0	14, 2, 0
D	12, 4, 0	12, 2, 0

C

Figure 4

In this game, player 1 chooses Rows, player 2 chooses Columns, and
player 3 chooses matrices; players 1 and 2 are long-lived while player 3 is
short-lived. The unique one-shot equilibrium is (U,L,A) with payoff (1,1,1).
The long-lived players can obtain a higher payoff if they induce player 3 to
choose C, which requires both long-run players to use mixed strategies. [Let
p = prob U; q = prob L, then player 3 chooses C if $(1-p)(1-q) \geq 1/10$ and $pq \geq$
1/100]. For example, if both long-run players use 50-50 randomization the
payoffs are (13,3,0). Call this strategy $\sigma = (\sigma_1, \sigma_2)$.

Now let us explain how to enforce (2,2) as an equilibrium payoff. It will be clear that the construction we develop is somewhat more general than the example; we do not give the general version because it does not lead to a complete characterization of the equilibrium payoffs. As in the proofs of Propositions 1 and 2, the strategies we construct depend on the history of the game only through a number of "state variables," with the current state determined by last period's state and last period's outcome through a (commonly known) transition rule. Let D and R be the "first" strategies, denoted $s_1(1)$, and $s_2(1)$, and let U and L be the second, $s_1(2)$ and $s_2(2)$.

Play begins in state 0. In this state, each player pays σ, which gives equal probability weight to his two actions. If player 1 plays action j, and player 2 action k, the next period state is (j,k). The payoffs when beginning play in state (j,k) are denoted $\alpha(k,k) = (\alpha_1(j,k), \alpha_2(j,k))$. In our construction, each player's continuation payoff will be independent of his opponents' last move, so that $\alpha_1(j,k) = \alpha_1(j)$ and $\alpha_2(j,k) = \alpha_2(k)$. Finally, in each state, the α's and the specified transition rule will be such that each player is indifferent between his pure strategies and thus is willing to randomize.

We will find it convenient to <u>first</u> define the α's, and <u>then</u> construct the associated strategies. Set the α's so that

(6)
$$\bar{v}_1 = (1-\delta)\, g_1(s_1(j),\, \sigma_2) + \delta\alpha_1(j)$$

$$\bar{v}_2 = (1-\delta)\, g_2(\sigma_1, s_2(k)) + \delta\alpha_2(k)\ .$$

In our example, $\alpha_1(1)$ solves $2 = (1-\delta)(12) + \delta\alpha_1(1)$, so $\alpha_1(1) = 12 - 10/\delta$. Similar computations yield $\alpha_1(2) = 14 - 12/\delta$, $\alpha_2(1) = 4 - 2/\delta$, and

$\alpha_2(2) = 2$.

If the observed play is $(s_1(1), s_2(1))$, next period's state is $(1,1)$, with payoffs $\alpha(1,1)$. Here play depends on a public randomization as follows. Choose a point w in the set P of payoffs attainable without private randomizations, and a probability $p \in (0,1)$, such that

$$(7) \qquad p(1-\delta)w + (1-p)\bar{v} = (1-\delta p)\alpha(1,1) \;.$$

(This is possible for δ sufficiently near to one. The general version of this construction imposes a requirement that guarantees (7) can be satisfied). With probability p, players play the strategies that yield w, and the state remains at $(1,1)$. With complementary probability, they play the mixed strategies (σ_1, σ_2). The continuation payoffs are <u>exactly</u> as at state 0. Thus, the payoff to player i of choosing strategy $x(j)$ is

$$(8) \quad p[(1-\delta)w_i + \delta\alpha_i(1)] + (1-p)[(1-\delta)g^i(s_i(j), a_{-i}) + \delta\alpha_i(j)] =$$
$$p[(1-\delta)w_i + \delta\alpha_i(1)] + (1-p)[g^i(s_i(1), a_{-i})] = \alpha_i(1) \;,$$

for all strategies $s(j) \in \text{supp}(a_i)$. Once again, if any long-run player chooses an action not in the support, play reverts to the static Nash equilibrium.

Now we must specify strategies at states other than $(1,1)$. The state $(2,2)$ occurs if the players chose their most preferred strategic (U,L). Choose a point w' in P and a $p' \in (0,1)$ such that

$$(9) \qquad p'(1-\delta)w' + (1-p')\bar{v} = (1-\delta p')\,\alpha(c_1, c_2)$$

(Once again this is possible for δ near to one.) And again, play depends on a public randomization, switching to w' with probability p' and otherwise following (a_1, a_2).

At state $(1,2)$, play switches to a point in P for one period with probability p to normalized payoffs out to $\alpha(1,2)$. With complementary probability, players once again play the mixed strategies σ and the same continuation payoffs are used. State $(2,1)$ is symmetric.

Now let us argue that the constructed strategies are an equilibrium for δ sufficiently large. First, if there are no deviations, the payoffs starting in state (j,k) are $\alpha(j,k)$. If player i deviates to an action outside of the support of a_i, or if either player deviates when the strategies say to play a pure strategy point, the deviation is detected, and play reverts to a static equilibrium. For δ near to one all of the α's exceed to static equilibrium payoff, and so for sufficiently large δ no player will choose such a deviation. And, by construction, players are indifferent between the actions in the support of their mixed strategy, if they plan to always conform in the future. Then by the principle of optimality, no arbitrary sequence of unilateral deviations is profitable.

In this appendix we consider discrete-parameter martingales $\{x_n, F_n\}$ n = 0,1,..., where $\{F_n\}$ is a filtration on an underlying probability space. We assume that $x_0 \equiv 0$.

Lemma A.1. Let $\{x_n, F_n\}$ be a martingale sequence with bounded increments. (That is, for some number B, $|x_n - x_{n-1}| \leq B$, almost surely.) Then $\lim_{n \to \infty} x_n/n = 0$ almost surely. A proof of this lemma can be found in Hall and Heyde (1980, page 36ff).

We also use the following standard adaptation of this strong law:

Lemma A.2 For $\{\xi_n, F_n\}$ as above, let $X_n = \min_{1 \leq n} x_i$. Then $\lim_{n \to \infty} X_n/n = 0$ almost surely.

Proof: Since $x_0 = 0$, $X_n \leq 0$ for all n. Fix a sample of the stochastic process. Since $X_n/n \leq 0$, we only have to show that lim inf $X_n/n = 0$. Suppose, instead, that n_i is a subsequence along which the limit is less than 0. For each n_i, there is $m_i \leq n_i$ with $x_{m_i} = X_{n_i}$, and thus $0 > X_{n_i}/n_i = x_{m_i}/n_i \geq x_{m_i}/m_i$. Hence, along the subsequence $\{m_i\}$, x_{m_i}/m_i violates the strong law, which can happen only on a null set.

Lemma A.3 Let $\{x_n, F_n\}$ be a supermartingale with bounded increments and with $x_0 \equiv 0$. Let $\{X_n\}$ be defined from $\{x_n\}$ as in lemma 2. Then $\lim_{n \to \infty} (x_n - X_n)/n = 0$ almost surely.

Proof: Since $x_n \geq X_n$, we only need to show that the limsup of the sequence is nonpositive. For n = 1, ..., let $\xi = x_n - x_{n-1}$, and let $\zeta = \xi_n - E(\xi_n | F_{n-1})$. Note that $\zeta_n \geq \xi_n$. Let $y_n = \Sigma_{i=1}^{n} \zeta_n$, and let $Y_n = \inf\{y_i; i = 1, ..., n\}$. Then, immediately, $\{y_n, F_n\}$ is a martingale sequence with bounded increments, and lemmas 1 and 2 tell us that $\lim y_n/n = \lim Y_n/n = 0$, and thus $\lim(y_n - Y_n)/n = 0$. We are done, therefore, once we show that $x_n - X_n \leq y_n - Y_n$ point wise. But this is easily done by induction. It is clearly true for n = 0 by convention. Assume it holds for n - 1; then since $\xi_n \leq \zeta_n$,

$$x_{n-1} - X_{n-1} + \xi_n \leq y_{n-1} - Y_{n-1} + \zeta_n, \text{ or}$$
$$x_n - X_{n-1} \leq y_n - Y_{n-1}.$$

If $X_n = X_{n-1}$, then since $Y_{n-1} \geq Y_n$, we are done. While if $X_n \neq X_{n-1}$, then X_n x_n, and $x_n - X_n \leq y_n - Y_n$. Q.E.D.

A symmetrical argument completes the proof, and we obtain:

Lemma A.4 : Let $\{x_n, F_n\}$ be a submartingale with bounded increments and $x_n \equiv 0$. Let $X_n = \max \{x_i | i = 1, ..., n\}$. Then $\lim_{n \to \infty}(X_n - x_n)/n = 0$ almost surely.

1. The required discount factor can depend on the payoffs to be attained.

2. We thank Ian Johnstone for pointing us to this result.

REFERENCES

Abreu, D. [1986] "On the Theory of Infinitely Repeated Games with
Discounting", mimeo.

Dybvig, P. and C. Spatt [1980] "Does it Pay to Maintain a Reputation?", mimeo.

Fudenberg, D. and D. Levine [1983] "Subgame-Perfect Equilibria of Finite and
Infinite Horizon Games", Journal of Economic Theory, 31, 227-256.

Fudenberg, D., and E. Maskin [1986] "The Folk Theorem in Repeated Games with
Discounting or With Incomplete Information," Econometrica, 54, 533-556.

_____ [1987a] "Discounted Repeated Games with One-Sided Moral Hazard",
Mimeo, Harvard University.

_____ [1987b] "Nash and Perfect Equilibria of Discounted Repeated Games,"
Mimeo.

_____ [1987c] "On the Dispensability of Public Randomizations in
Discounted Repeated Games," mimeo.

Hall, P. and C.C. Heyde, [1980] Martingale Theory and Its Applications,
Academic Press, New York.

Kreps, D. [1984] "Corporate Culture." Mimeo, Stanford Business School.

Radner, R. [1986] "Repeated Partnership Games with Imperfect Monitoring and no
Discounting." Review of Economic Studies 53, 43-58.

Radner, R., R. Myerson, and E. Maskin [1986] "An Example of a Repeated
Partnership Game with Discounting and With Uniformly Inefficient
Equilibria." Review of Economic Studies 53, 59-70.

Selten, R. [1977] "The Chain-Store Paradox.", Theory and Decision, 9, 127-159.

Shapiro, C. [1982] "Consumer Information, Product Quality, and Seller
Reputation," Bell Journal of Economics, 13, 20-35.

Simon, H. [1951] "A Formal Theory of the Employment Relationship,"
Econometrica, 19, 293-305.

Date Due

Lib-26-67

3 9030 005 356 594